

CHAPTER 16

VALIDITY

As noted in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1985, p. 9), “validity is the most important consideration in test evaluation.” Validity refers to whether specific inferences made from test scores are appropriate, meaningful, and useful. There are several types of validity-related evidence that can be used to support appropriate, meaningful, and useful inferences based on test scores.

CONTENT-RELATED EVIDENCE

As noted in the *Standards* (p. 10), evidence of test validity begins with test development and continues throughout the entire testing process. Chapters 2 through 9 provide evidence regarding the alignment between the content of the MEA and Maine’s *Learning Results*.

Further internal validity evidence is exhibited by correlations computed for subscores of the different tests¹. For each subject in each grade, with the exception of Reading, two subscores were computed based on whether items pertain to content or application. For Reading, three subscores were computed – process, comprehension, and research. Within each grade the correlations of subscores within and across subjects were computed and presented in Tables 16-1a to 16-3a. The average correlations within and across subject were also computed and presented in Tables 16-1b to 16-3b.

¹ The analyses discussed here were performed only on common item tests.

Table 16-1a Subscore Correlations: Grade 4										
Subject	Subscore	Reading			Math		Science		Soc.	
		Process	Comprehension	Research	Content	Application	Content	Application	Content	Application
Reading	Process Comprehension Research	1								
		0.51	1							
		0.45	0.65	1						
Mathematics	Content Application	0.34	0.46	0.46	1					
		0.43	0.61	0.60	0.61	1				
Science & Technology	Content Application	0.43	0.55	0.52	0.45	0.58	1			
		0.36	0.48	0.43	0.36	0.48	0.55	1		
Social Studies	Content Application	0.41	0.56	0.53	0.44	0.58	0.51	0.42	1	
		0.35	0.53	0.50	0.38	0.52	0.43	0.36	0.52	1
Note: Process = Reading Process and Language (A & C) Comprehension = Reading Comprehension (B & D) Research = Research Skills (H)										

Table 16-1b Average Correlations: Grade 4	
Mathematics	0.61
Science	0.55
Reading	0.54
Social Studies	0.52
Reading & Social Studies	0.48
Reading & Math	0.48
Math & Social Studies	0.48
Math & Science	0.47
Reading & Science	0.46
Social Studies & Science	0.43

Table 16-2a Subscore Correlations: Grade 8										
Subject Subscore		Reading			Math		Science		Soc.	
		Process	Comprehension	Research	Content	Application	Content	Application	Content	Application
Reading	Process Comprehension Research	1								
		0.57	1							
		0.48	0.63	1						
Mathematics	Content Application	0.45	0.57	0.57	1					
		0.46	0.57	0.57	0.76	1				
Science & Technology	Content Application	0.45	0.53	0.50	0.61	0.57	1			
		0.49	0.60	0.57	0.57	0.60	0.58	1		
Social Studies	Content Application	0.49	0.60	0.56	0.60	0.59	0.60	0.59	1	
		0.48	0.61	0.58	0.54	0.56	0.52	0.61	0.63	1
Note: Process = Reading Process and Language (A & C) Comprehension = Reading Comprehension (B & D) Research = Research Skills (H)										

Table 16-2b Average Correlations: Grade 8	
Mathematics	0.76
Science	0.63
Reading	0.58
Social Studies	0.56
Reading & Social Studies	0.59
Reading & Math	0.58
Math & Social Studies	0.57
Math & Science	0.55
Reading & Science	0.53
Social Studies & Science	0.52

Table 16-3a Subscore Correlations: Grade 11										
Subject	Subscore	Reading			Math		Science		Soc.	
		Process	Comprehension	Research	Content	Application	Content	Application	Content	Application
Reading	Process Comprehension Research	1								
		0.65	1							
		0.56	0.66	1						
Mathematics	Content Application	0.41	0.48	0.52	1					
		0.48	0.57	0.60	0.73	1				
Science & Technology	Content Application	0.45	0.54	0.53	0.56	0.66	1			
		0.49	0.58	0.58	0.53	0.63	0.63	1		
Social Studies	Content Application	0.34	0.42	0.40	0.41	0.47	0.47	0.45	1	
		0.51	0.60	0.59	0.54	0.63	0.61	0.63	0.51	1
Note: Process = Reading Process and Language (A & C) Comprehension = Reading Comprehension (B & D) Research = Research Skills (H)										

Table 16-3b Average Correlations: Grade 11	
Mathematics	0.73
Science	0.63
Reading	0.62
Social Studies	0.51
Reading & Social Studies	0.60
Reading & Math	0.54
Math & Social Studies	0.53
Math & Science	0.51
Reading & Science	0.51
Social Studies & Science	0.48

In grade 4, Mathematics had the highest average correlation (.63) out of the four subject areas. The other three subjects were fairly comparable in terms of average correlation: science (.55), reading (.54) and social studies (.52). The average correlation for each subject (.52 to .61) was greater than the average correlation between each pair of subjects (.43 to .48). The correlations between the subjects were similar. Five out of six pairs showed virtually no difference, ranging from .46 to .48. Only the correlation between social studies & science (.43) was outside this range.

It appears that mathematics, by its nature is the most homogeneous construct (i.e., compared to the other subjects). This phenomenon is consistent with other assessment. The comparable correlations within the other three subjects (i.e., science, reading, and social studies) reflect the multi-faceted nature of each subject. It is not surprising that these numbers are lower than the correlation for mathematics. As expected, the average correlation within each subject was greater than the average correlation between each pair of subjects. One additional observation is worth noting. The between group correlations were all about the same (.43) to (.48). This suggests that students at this level have not yet developed “specialties” in one area over another. It suggests that students, who read well, do well on social studies, and math, and science too.

Just like in grade 4, grade 8 Mathematics had the highest average correlation (.76) out of the four subject areas. The average correlation for reading (.56) was lower than social studies (.63) and science (.58), but was higher than the correlation for grade 4 reading (.52). The correlations between subjects were similar ranging from .52 to .59. The correlation between math & science was highest (.59). The correlation between reading & math (.53) and reading & science (.52) were lowest.

As was the case in grade four, mathematics yielded the highest within subject correlation. Although there was more variance among the numbers (i.e., as compared to grade four), the correlations within social studies, science, and reading were fairly comparable. Between subjects, it is not surprising that the correlation between math & science was highest because one might expect students who are mathematically inclined to be slightly more interested in science. Somewhat surprising was the strong relationship between social studies & science (.58) and social studies & math (.57). However, if math & science were highly correlated, and science & social studies were highly correlated, then (of course) math & social studies would be highly correlated. The results indicated that reading & math and reading & science yielded the lowest correlations. Reading & math are probably the two subjects that differ most in content, so it's not surprising they have the smallest correlation. From these results it becomes apparent that science is more closely tied to math ability (.59) than reading

ability (.52). Finally, the large correlations found between some subjects (e.g., math & science, social studies & science, and math & social studies), as opposed to within some subjects (e.g., science and reading) appears to be a product of the high correlation within math and the low correlation within reading. These patterns seem to make sense. The only surprise was that social studies was more highly correlated with math than with reading.

For grade 11, Mathematics also had the highest average correlation (.73) out of the four subject areas. The average correlation for science (.63) and reading (.62) were similar. Social studies had the lowest average correlation (.51). The correlation between math & science was highest (.60). The correlations between social studies & science (.54) and reading & science (.53) were similar. The correlations between social reading & math (.51) and math & social studies (.51) were the same. Surprisingly, the correlation between reading & social studies was lowest (.48).

As was the case in grade four and eight, mathematics yielded the highest within subject correlation. These results suggest across all three grades, mathematics is the most homogeneous construct or subject. In grade 11, the within subject correlation for social studies was much lower than reading, science and mathematics. This may be explained by the tendency of high school students to be less motivated to perform well on this test than on others such as math, reading or science.

In grades 8 and 11, the relationship between mathematics & science was the strongest. In high school, when science means chemistry and physics, mathematics ability is crucial to success. It makes sense that students who are mathematically inclined are more likely to do well in science. The correlations between the other subjects are generally similar. There is slightly less variation in grade 11 than in grade 8. The one exception is the relationship between reading & social studies (.48). One might expect social studies to correlate more highly with reading than with science. These results suggest that: a) social studies is the most diverse subject (e.g., measuring civics and government, history, geography, and economics), and b) the relationship between social studies ability and math ability in grade 11 is stronger than the relationship between social studies ability and reading ability.

EXTERNAL EVIDENCE

External validity of the MEA is conveyed by the relationship of test scores and situational variables such as school transience, course-taking pattern, attitude towards subject matter, and self-image. These situational variables were all based on student questionnaire data collected during the

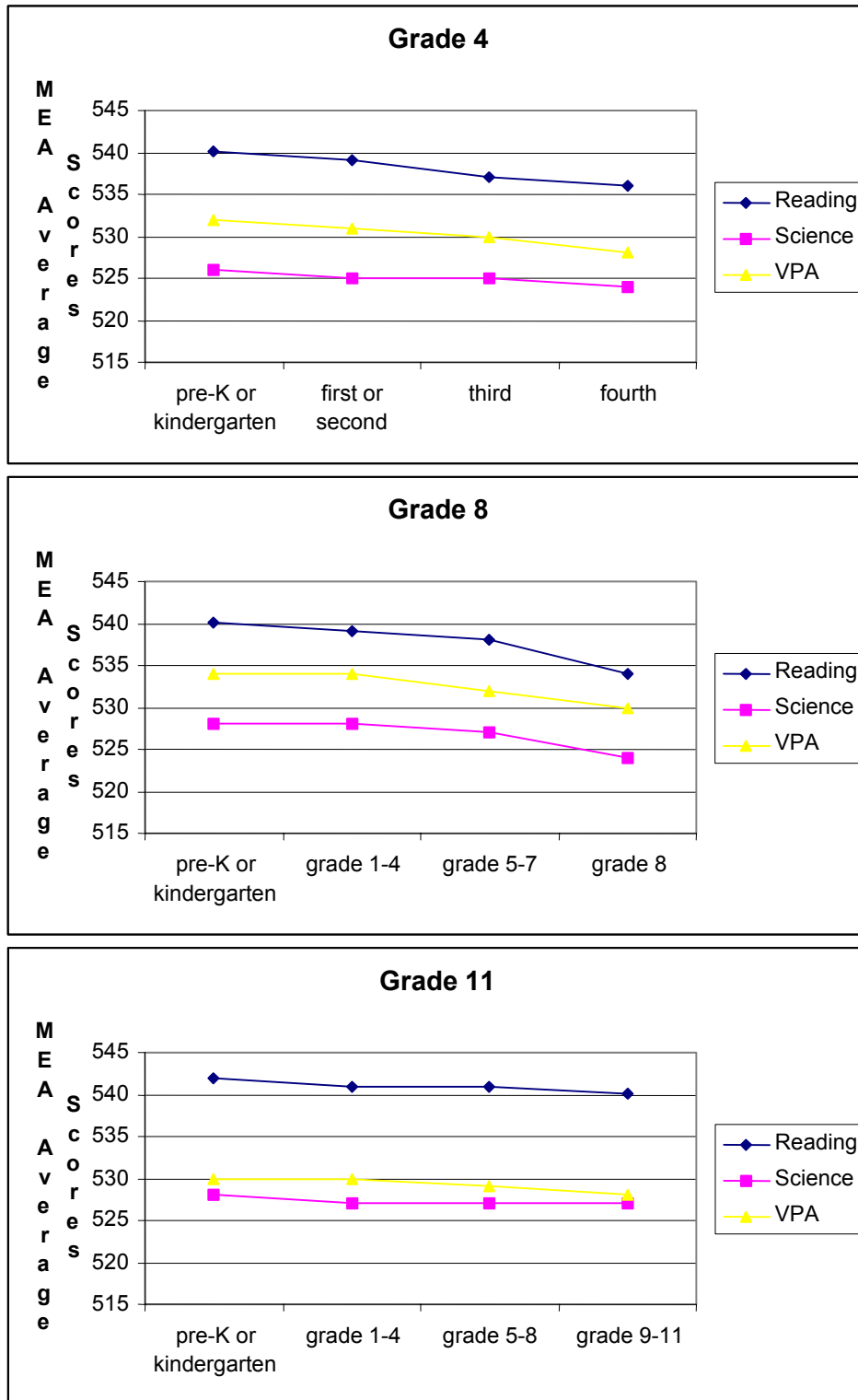
administration of the MEA. Note that not all the questionnaire items referred to in the following subsections were asked regarding all of the subjects assessed by the MEA. Note also that no inferential statistics are included. However, because the numbers of students are large enough, differences in average scores could be shown to be statistically significant.

SCHOOL TRANSIENCE

This is an evaluation of how time in a single school is related to test scores. Students were asked, “In what grade did you start coming to school in this school district?” Medsker (1998) found that typically, students who change schools often do not perform as well as students who regularly attend a single school or school system. Charts in Figure 16-1 clearly indicate that students who spent more time in a single school tended to have higher test scores in reading, science, and visual and performing arts.

Figure 16-1
School Transience and MEA Scores

Question: In what grade did you start coming to school in this school district?

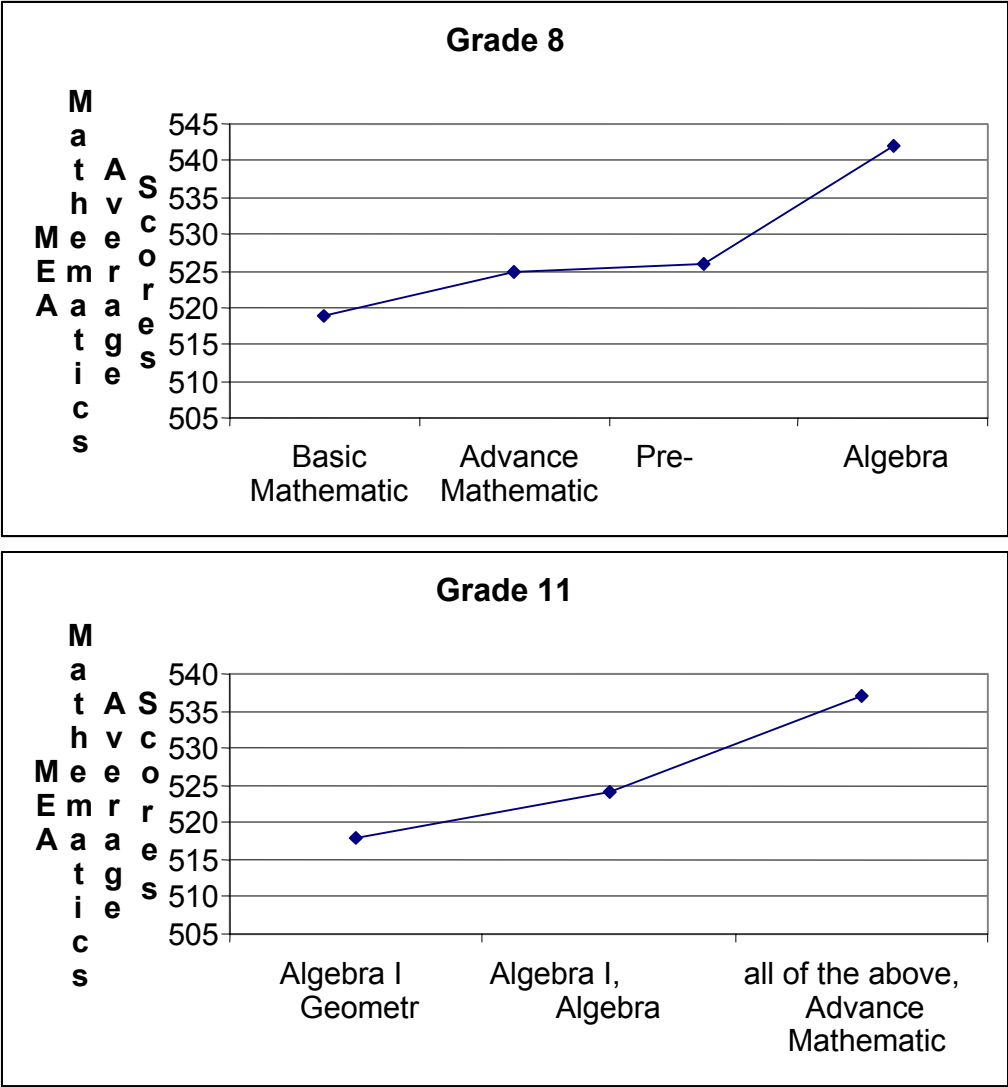


COURSE-TAKING PATTERN

Grades 8 and 11 examinees were asked questions related to their course-taking patterns in mathematics. Eighth graders were asked, “What best describes the mathematics class you are taking in the eighth grade?” and eleventh graders were asked, “What best describes the mathematics courses you will complete before you graduate?” Charts in Figure 16-2 both show that the higher level mathematics courses are associated with higher MEA mathematics scores.

Figure 16-2
MEA Mathematics Scores and Course-Taking Patterns

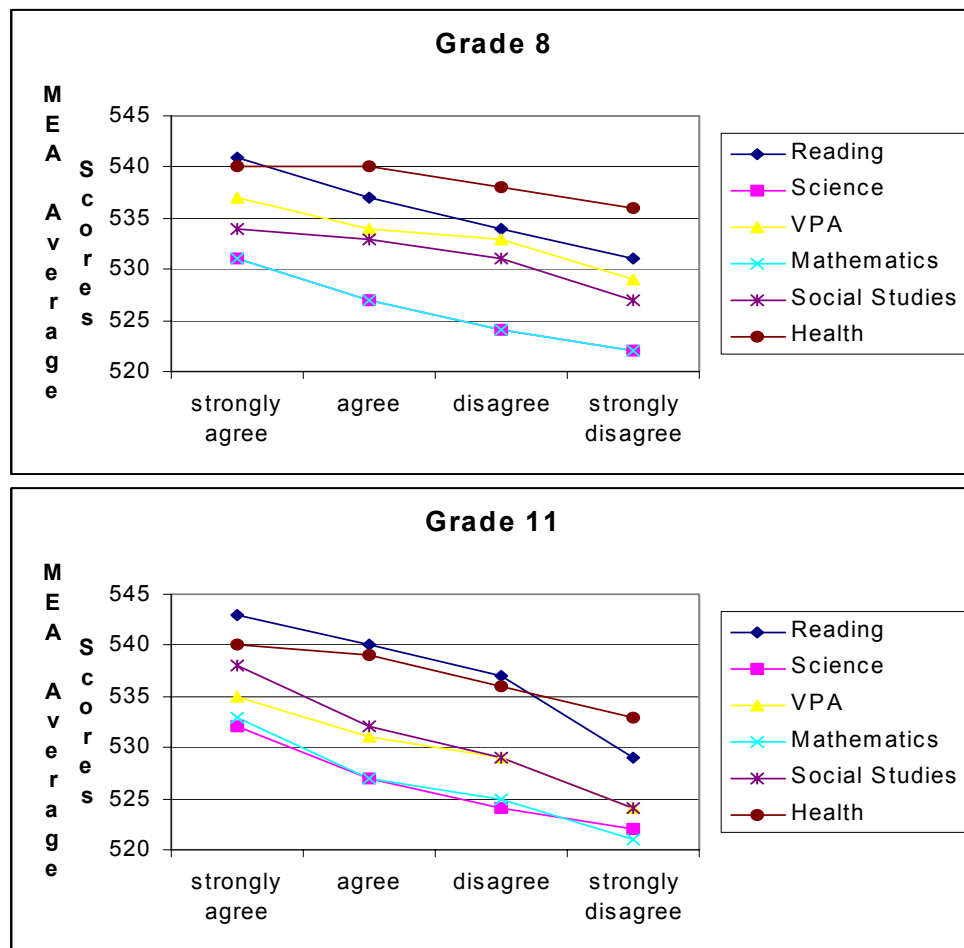
Grade 8 Question: What best describes the mathematics class you are taking in the eighth grade?
Grade 11 Question: What best describes the mathematics courses you will complete before you graduate?



Questionnaire items related to examinees' attitudes toward different subjects tested in the MEA were administered to eighth and eleventh graders. For reading, mathematics, science, social studies, and visual and performing arts, students were asked how they feel about the statement, "My knowledge of [subject] will be useful to me in my future work." For health, students were asked how they feel about the statement, "My knowledge about health education will be helpful to me as an adult." Charts in Figures 16-3 both show that students degree of agreement with statements that indicate their attitudes toward the subjects tested in the MEA are related positively with MEA scores.

Figure 16-3
Attitude Towards Subject Matters and MEA Scores

Question: My knowledge of [subject] will be useful to me [in my future work/as an adult].



SELF-IMAGE

All students who participate in the MEA were asked, “What best describes how you rate yourself as a student?” Figure 16-4 indicates that there is a positive relationship between students’ self-image and their MEA scores in mathematics, social studies, and health.

Figure 16-4
Self-Image and MEA Scores

Question: What best describes how you rate yourself as a student?

